TU Clausthal

Clausthal University of Technology

Markov Temporal Logic

Wojciech Jamroga

IfI Technical Report Series **IfI-07-11**

Department of Informatics
Clausthal University of Technology

Impressum

Publisher: Institut für Informatik, Technische Universität Clausthal
Julius-Albert Str. 4, 38678 Clausthal-Zellerfeld, Germany
Editor of the series: Jürgen Dix
Technical editor: Wojciech Jamroga
Contact: wjamroga@in.tu-clausthal.de
URL: http://www.in.tu-clausthal.de/forschung/technical-reports/
ISSN: 1860-8477

The IfI Review Board

Markov Temporal Logic

Wojciech Jamroga

Department of Informatics, Clausthal University of Technology
Julius Albert Str. 4, D-38678 Clausthal Germany
wjamroga@in.tu-clausthal.de

Abstract

Most models of agents and multi-agent systems include information about possible states of the system (that defines relations between states and and their external characteristics), and information about relationships between states. *Qualitative* models of this kind assign no numerical measures to these relationships. At the same time, *quantitative models* assume that the relationships are measurable, and provide numerical information about the degrees of relations. In this paper, we explore the analogies between some qualitative and quantitative models of agents/processes, especially those between transition systems and Markovian models. Typical analysis of Markovian models of processes refers only to the expected utility that can be obtained by the process. On the other hand, modal logic offers a systematic approach to describing phenomena by combining various modal operators. Here, we try to exploit linguistic features, offered by propositional modal logic, for analysis of Markov chains and Markov decision processes. To this end, we propose *Markov temporal logic* MTL– a multi-valued logic that extends the branching time logic CTL*.

1 Introduction

There are many different models of agents and multi-agent systems; however, most of them follow a similar pattern. First of all, they include information about possible situations (states of the system) that defines relations between states and and their external characteristics (essentially, "facts of life" that hold in these states). Second, they provide information about relationships between states (e.g, possible transitions between states).

Models that share this structure can be, roughly speaking, divided into two classes. *Qualitative models* provide no numerical measures for these relationships. Examples include automata, labeled and unlabeled transition systems, epistemic models (and combinations of these), flowcharts, data and/or control flow diagrams etc. They are widely used as basic models of computational systems, in semantics of programming languages (including agent-oriented languages), and in specification and verification of systems. Qualitative models seem especially suited for domains in which quantitative information cannot be reliably obtained nor assumed. In agent systems they are also used to model situations in which the goal of an agent (group of agents, the whole system) is not to maximize a measurable output, but rather to achieve a state that matches certain characteristics (specified e.g. by means of a logical formula).

Quantitative models assume that relationships are measurable, and provide numerical information about the degrees of relations. For transition relations between states, the degrees are usually given in the form of probabilities. For "qualities" of particular states, one often talks about *rewards* or *utilities*. Among other things, this allows to construct a ranking of states and/or actions, and support decision making. Quantitative representations are used in stochastic modeling (Markov chains), decision theory and reinforcement learning (Markov decision processes), game theory (strategic and extensive game forms) etc. The relationship between models and concepts of game theory and modal logic has been already studied in several places [21, 5, 36]. In this paper, we explore the analogies between transition systems and Markovian models in order to provide a more expressive language for reasoning about, and specification of agents in stochastic environments.

Analysis of quantitative process models is usually based on the notion of expected reward. However, there are other meaningful properties of a process that may be interesting to study – for example, the minimal and maximal reward that can be obtained. Of course, one can address these (and many more) characteristics in the general mathematical language of decision theory and/or game theory. Still, these theories use expressions of higher order logic, which goes too far in many cases. Propositional modal logic provides an intuitive language in which many phenomena can be described and studied in a systematic way. On one hand, the language is limited so that it enforces self-discipline when using it; also, the complexity of related computational problems is relatively low. On the other hand, it is expressive enough to enable specification of many important properties. In this paper, we propose to use the same methodology in order to study quantitative properties of systems and processes. Apart from the expected cumulative reward, there are other features of Markov chains and Markov decision processes which might be interesting. *Markov temporal logic* for Markov chains, introduced in Section 4, is our first step in this direction. We also briefly consider two extensions of the logic: first, for Markov decision processes (where a single decision maker is present); next, for multi-agent Markov decision processes (in which many agents can play simultaneously).

1.1 Related Work

The related work includes research on multi-valued logics, especially fuzzy logics [39, 18, 25], probabilistic logics [31, 32], and multi-valued modal logics [16, 13, 26]. Of the latter, [26] is particularly relevant, as it defines a multi-valued version of the branching-time logic CTL*, with propositions and accessibility relations taking values from a finite quasi-Boolean algebra. Still, the approach of [26] is too abstract to give an account of quantitative analysis of processes (e.g., by operators that compute the expected and/or average truth value along a given path).

Logics of probability [4, 35, 19] are also related to the phenomena we study here. Important examples of such logics are two probabilistic variants of CTL: PCTL [20] for real time, and pCTL* [3] for discrete time; both allow to express probability bounds for a specified behavior. However, logics of probability do not use the machinery of multi-valued logics. More importantly, like probabilistic logics, they focus on the probabilities of events (e.g., behaviors), and it is often hard to attribute an intuitive meaning to combinations (or patterns) of different probability values. In contrast, we will argue in Section 2.3 that combining utilities has a very natural commonsense interpretation.

Our work comes very close to [10, 11], where the "Discounted CTL" (DCTL) is proposed. In fact, our Markov temporal logic directly extends the ideas behind DCTL; a more detailed comparison is presented in Section 6. The variant of multi-valued CTL from [27], where the domain of truth values can be any c-semiring (rather than simply the interval $[0, 1]$ of real numbers), is also relevant. While it does not address quantitative analysis of processes directly, the choice of c-semirings makes such analysis possible (at least in principle). It may be interesting to consider a similar generalization of our framework in the future.

2 Looking for Analogies: the Quantitative and Qualitative Tradition

We begin with drawing some analogies between the quantitative and qualitative approaches to computational systems. In particular, we are interested in exploiting the similarities between Markovian models of processes and transition systems.

2.1 Quantitative vs. Qualitative Models of Processes

The simplest Markovian models are Markov chains [29, 24, 17]. A *Markov chain* (MC) is a discrete-time stochastic process with the Markov property. That is, it consists of a countable set of states, and a probabilistic transition

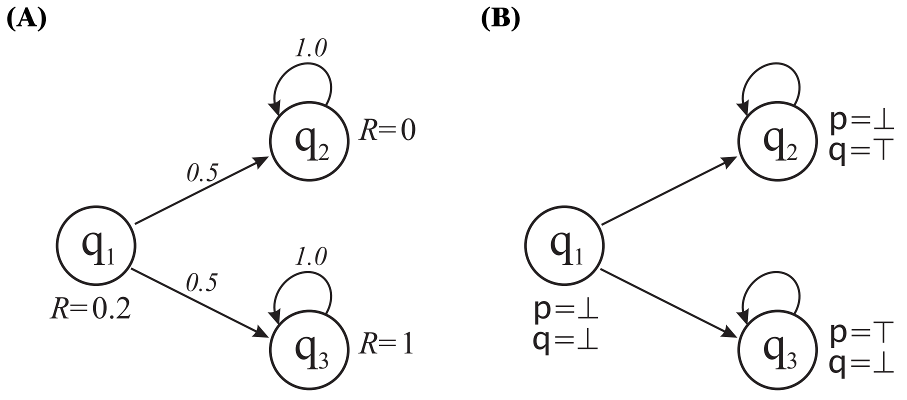

Figure 1: (A) Markov chain. (B) Unlabeled transition system

relation between states. The Markov property requires that the next state of the system depends only on the present state and possibly the present action(s), but it does not directly depend on the past states of the system. Reward/utility values are usually missing in Markov chains, but we include them for compatibility with Markov decision processes. A formal definition is given in Section 3.2.

An example Markov chain is depicted in Figure 1, together with an unlabeled transition system. It is easy to see the similarities. First, states in the Markov chain are assigned real reward values R, and states in the transition system are assigned truth values of atomic propositions $\mathsf{p}, \mathsf{q}, \ldots$. Moreover, both kinds of structures include a set of states and a (single) binary transition relation on states; however, in the MC, tuples of the relation are annotated with transition probabilities.

Markov decision processes [7, 6, 23, 33] can be seen as an extension of Markov chains, where several actions are available in each state. A *Markov decision process* (MDP) models a decision-making agent in a stochastic environment. We observe that Markov decision processes are very much like labeled transition systems. In both cases, the action-transition structure can be modeled by a number of binary relations on states (one relation per action), although the elements of relations in MDP are annotated with probability values (cf. Figure 2). Note that, by fixing the agent's policy in advance, we "instantiate" a Markov decision process to a Markov chain in the same way as labeled transition systems "instantiate" to unlabeled ones. For instance, if we assume that the agent always chooses action α, then the Markov decision process in Figure 2A reduces to the Markov chain from Figure 1A, and the labeled transition system in Figure 2B reduces to the unlabeled transition system from Figure 1B.

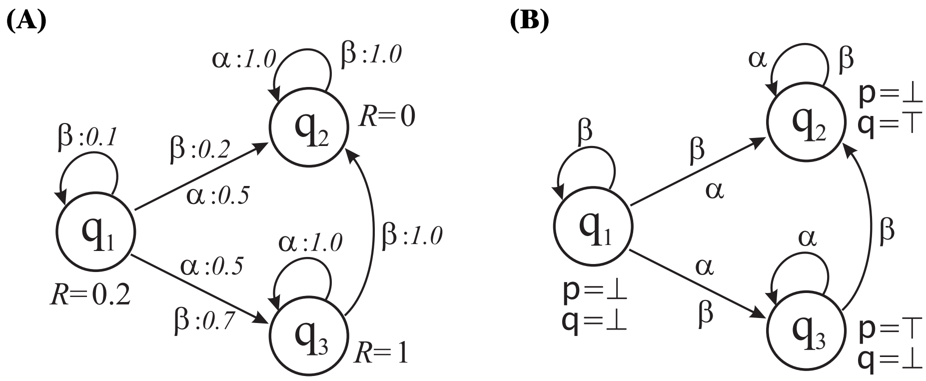

Figure 2: (A) Markov decision process. (B) Labeled transition system

Markov decision process have been "upgraded" to the multi-agent case in several ways (and, in fact, there are also various extensions of transition systems for MAS – depending on whether we assume synchrony or asynchrony, perfect or imperfect information, turn-based or possibly simultaneous play etc.). Here, we only observe the similarity between multi-agent Markov decision processes (MMDP) from [8] and concurrent game structures from [2]. In both cases, the "upgrade" has been done by defining a set of agents, and labeling transitions with *tuples* of actions, one action per agent. This corresponds to all agents acting simultaneously: the transition to the next state is a result of the combination of actions being played now.

2.2 Quantitative vs. Qualitative Descriptions

The tradition of decision theory and reinforcement learning puts forward the quantitative notion of expected utility which represents the average of "what we can get" for all possible executions of the process. At the same time, logical approaches are usually concerned with "limit properties" like the existence of an execution that displays a specific temporal pattern. In that case, we can state that such an execution exists, or that the pattern is universally displayed by all executions, but we cannot address the middle ground between these two extremes (e.g., what is the ratio of such executions etc.). In consequence, logical frameworks are not very well suited to coping with process models that involve probabilities: the existence of a particular kind of execution may be of little interest if this kind of execution is unlikely to happen; on the other hand, if all the possible behaviors of a real-life process satisfy some property, then the property is usually trivial. It does not mean, however, that these "limit properties" are irrelevant: in some cases we do want to e.g. make sure that there is no path violating an important security property. The point we are trying to make in this paper is that *both* kinds of properties are interesting and worth using to describe processes.

One of the nicer features of temporal logics – especially branching-time logics like CTL and CTL* – is that they offer a systematic approach in which properties of particular paths (executions) are distinguished from the properties of sets of paths (e.g., the set of all executions of a process). The first kind of properties is facilitated by *temporal operators* like "always" ($\Box$), "eventually" ($\Diamond$), "next" ($\bigcirc$) etc. The second kind is based on *path quantifiers* like "for all paths" (A) and "there is a path" (E). Both kinds of operators can be combined: e.g., $\mathsf{E}\Box\mathsf{safe}$ says "there is a path such that the system is always in a safe state". The same approach can be employed within the quantitative framework. For instance, besides the expected value of cumulative future reward, we can ask of the maximal (or minimal) cumulative reward. Or, we might be concerned with the expected value of minimal guaranteed reward etc.

The above examples show how "limit properties" (captured by operators of temporal logic) can be combined with "mean properties" of a process (expected value, cumulative or average reward) in a meaningful way. We propose a precise semantics for such combinations (and a semantics of interplay between qualitative and quantitative properties) in Section 4. But first, we observe that operators of classical and temporal logic have a very intuitive quantitative interpretation as maximizers and minimizers of truth values.

2.3 Logical Operators as Minimizers and Maximizers

Note that – when truth values represent the utility of an agent – temporal operators "sometime" and "always" have a very natural interpretation. First, "sometime p" ($\Diamond p$) can be rephrased as "p is achievable in the future". Under the assumption that agents want to obtain as much utility as possible, it is natural to view the operator as maximizing the utility value along a given temporal path. Similarly, "always p" ($\Box p$) can be rephrased as "p is guaranteed from now on". In other words, $\Box p$ asks for the minimal value of p on the path. On a more general level, every universal quantifier is essentially a minimizer of truth values, while existential quantifiers can be seen as maximizers. Thus, $\mathsf{A}\gamma$ ("for all paths γ") minimizes the utility specified by γ across all paths that can occur, etc. Also, conjunction and disjunction can be seen as a minimizer and a maximizer: $\varphi \vee \psi$ reads easily as "the utility that can be achieved through φ or ψ", while $\varphi \wedge \psi$ reads as "utility guaranteed by both φ and ψ".

Of course, the idea of defining semantics of conjunction and disjunction through functions $\min$ and $\max$, respectively, is not new: the same semantic approach is used e.g. in fuzzy logic [39, 18, 25]. Also, interpreting quantifiers as outcome maximization/minimization operators, can be traced back to the game semantics of classical logic [22, 28].

3 Basic Models: Markov Chains and Markov Decision Processes

Markov chains have been proposed to represent and study properties of processes in which transitions can be described in terms of probabilities. In particular, the processes are required to satisfy the *Markov property* that the probability distribution for transitions in state q depends only on the state (and thus it is independent from the past transitions that led to q). Thus, the behavior of the system can be described by a collection of conditional probabilities $\tau(q, q') = pr(Next = q' \mid Now = q)$.

Markov chains are often used for generation of semi-random sequences of words, symbols or events (algorithms generating spam messages are a good example here). For these applications, states of a system (chain) play mostly a technical role, as we are mainly after the events being generated. However, Markov chains can be also used to model and analyze existing processes (especially as parts of *Markov decision processes*, perhaps the most popular models of reinforcement learning). In that case, we are usually interested in properties of the states: either qualitative (i.e., some facts being true or false in different states of the process) or quantitative (representing utilities or rewards that the process is expected to yield in particular states). Even more importantly, we are interested in how these (qualitative or quantitative) properties accumulate as the system progresses in time.

3.1 Domain

A *domain* $D = \langle U, \top, \bot, ^- \rangle$ consists of: (1) a set $U \subseteq \mathbb{R}$ of *utility values* (or simply *utilities*); (2) special values $\top, \bot$ standing for the logical truth and falsity, respectively; $\hat{U} = U \cup \{\top, \bot\}$ will be called the *extended utility set*; and, finally, (3) a complement function $^- : \hat{U} \to \hat{U}$. A domain should satisfy the following conditions:

1. $U \subseteq \mathbb{R}$;
2. The operations of addition and multiplication have their typical properties on $\hat{U}$, and $\hat{U}$ is closed under averaging, i.e., for every probability distribution P over $\hat{U}$ (discrete or continuous), $\sum_{u \in \hat{U}} u\, P(u) \in \hat{U}$;
3. U is closed under complement: if $u \in U$ then $\overline{u} \in U$;
4. Complement reverts the classical truth values: $\overline{\top} = \bot$ and $\overline{\bot} = \top$;
5. $\top \geq 0$;

6. $\bot \le u$ and $\top \ge u$ for all $u \in \hat{U}$;[1]
7. The complement is quasi-boolean wrt $\max, \min$, i.e., for every $u_1, u_2, u \in \hat{U}$: $\overline{\max(u_1, u_2)} = \min(\overline{u_1}, \overline{u_2})$, $\overline{\min(u_1, u_2)} = \max(\overline{u_1}, \overline{u_2})$, $\overline{u_1} \le \overline{u_2}$ iff $u_2 \le u_1$, and $\overline{\overline{u}} = u$.

In the rest of the paper, we will assume that $U = [0, 1], \top = 1, \bot = 0, \overline{u} = 1 - u$ (unless explicitly stated otherwise). This closely resembles the setting in [10, 11]. Admittedly, using 0 and 1 to represent "false" and "true" has a long tradition in logic; there is also a tradition of using values between 0 and 1 in multi-valued logics.

Another meaningful example of a domain is $U = (-1, 1)$, $\top = 1$, $\bot = -1$, and $\overline{u} = -u$. It has some advantages over $[0, 1]$ since it allows to explicitly represent both gains and losses rather than only bigger and smaller gains. Even more importantly, $U = (-1, 1)$ allows to distinguish the "zero gain" situation from the classical truth value of "false". This way, it is possible to distinguish between a specification being (qualitatively) false and a specification yielding a (quantitative) utility of 0. In more general terms, we have $\top \notin U$ and $\bot \notin U$, which makes a clear distinction between quantitative and qualitative semantic values.

Perhaps the best solution in pragmatic terms would be to impose no restrictions on rewards, i.e. to assume $U = \mathbb{R}$. In such case, the classical truth values can be thought of as $-\infty, +\infty$. However, this leads to some problems with defining the value of numerical expressions, and we leave proper exploration of this possibility for future work.

3.2 Markov Chains

Typically, a Markov chain is a directed graph with probabilistic transition relation. In our definition, we include also a device for assigning states with utilities and/or propositional values. This is done through *utility fluents* which generalize atomic propositions from modal logic.

Definition 1 (Markov chain) *A* Markov chain *over domain $D = \langle U, \top, \bot, ^{-} \rangle$, and a set of utility fluents* Π *is a tuple $M = \langle St, \tau, \pi \rangle$, where:*

- *St is a set of* states *(we will assume that the set is finite and nonempty throughout the rest of the paper);*
- *$\tau : St \times St \to [0, 1]$ is a* stochastic transition relation *that assigns each pair of states q_1, q_2 with a probability $\tau(q_1, q_2)$ that, if the system is in q_1, it will change its state to q_2 in the next moment. For every $q_1 \in St$, $\tau(q_1, \cdot)$ is assumed to be a probability distribution, i.e. $\sum_{q_2 \in St} \tau(q_1, q_2) = 1$.*

[1] Note that this implies that $\max(u, \top) = \top$, $\min(u, \top) = u$, $\min(u, \bot) = \bot$, and $\max(u, \bot) = u$ for all $u \in \hat{U}$.

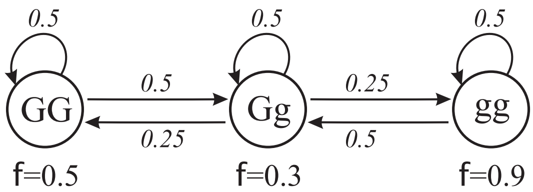

Figure 3: Markov chain for the gene model

By abuse of notation, we will sometimes write $\tau(q)$ to denote the set of states accessible in one step from q, i.e. $\{q' \mid \tau(q,q') > 0\}$.

- $\pi : \Pi \times St \to \hat{U}$ *is a* valuation of utility fluents.

Consider the following extension of the "gene model" example from [17].

Example 1 (Gene model) *Suppose that a trait in animals of a particular species is governed by a pair of genes, each of whom may be of type G or g. Very often the GG and Gg types are indistinguishable in appearance; we say that type G dominates type g. Thus, an individual may have the* dominant *combination GG,* recessive *combination gg, or* hybrid *combination Gg (which is genetically the same as gG).*

Mating of two animals produces an offspring that inherits one gene of the pair from each parent, and the basic assumption of genetics is that these genes are selected at random, independently of each other. Suppose that we breed animals by starting with an individual of known genetic character and mate it with a hybrid. We assume that there is at least one offspring. Then, at each round, a random offspring is chosen and mated with a hybrid, and so on. Suppose also that a statistical study of survival produced the following fitness function for individuals of the species (in relation to genotype): $f(GG) = 0.5$, $f(Gg) = 0.3$, and $f(gg) = 0.9$ – i.e., the individuals with recessive genes are the fittest, and hybrids are the least fit of all. A Markov chain that models the process is shown in Figure 3.

A *run* in Markov chain M is an infinite sequence of states $q_0q_1\ldots$ such that each q_{i+1} can follow q_i with a non-zero probability, i.e., for every $i = 0, 1, \ldots$ we have $\tau(q_i, q_{i+1}) > 0$. We denote the set of all runs in M by $\mathcal{R}_M$. The set of runs starting from state q is denoted by $\mathcal{R}_M(q)$.[2] Let $\lambda = q_0q_1...$ be a run and $i \in \mathbb{N}_0$. Then: $\lambda[i] = q_i$ denotes the ith position in λ; $\lambda[i..j] = q_i \ldots q_j$ denotes the subpath of λ from position i to j; and $\lambda[i..\infty] = q_iq_{i+1}\ldots$ denotes the infinite subpath of λ from position i on.

Finite prefixes of runs are called *histories*. $\mathcal{H}_M = \{h \mid h = \lambda[0..i]$ for some $\lambda, i\}$ denotes the set of all histories in M; and the set of histories starting from q by $\mathcal{H}_M(q)$. $\mathcal{H}^k_M(q)$ restricts the set further to the histories of length k. Note

[2] If the model is clear from the context, the subscripts will be omitted.

that each history h can be uniquely identified with the set of runs that "complete" it. By a slight abuse of notation, we will also use h to denote the set, and $\mathcal{H}_M(q)$ to denote all such subsets of $\mathcal{R}_M(q)$. Finally, by $\lambda(h)$ we denote an arbitrary infinite continuation of h (e.g., the run completing h which is minimal wrt to alphabetical ordering of runs).

3.3 Markov Decision Processes

Markov decision processes extend Markov chains with an explicit action structure: transitions are now connected to actions that generate them.

Definition 2 (Markov decision process) *A* Markov decision process *over domain* $D = \langle U, \top, \bot, ^- \rangle$*, and a set of utility fluents* Π *is a tuple* $\mathcal{M} = \langle St, Act, \tau, \pi \rangle$*, where:* St, π *are like in a Markov chain,* Act *is a nonempty finite set of actions, and* $\tau : St \times Act \times St \to [0, 1]$ *is a stochastic transition relation;* $\tau(q_1, \alpha, q_2)$ *defines the probability that, if the system is in* q_1 *and the agent executes* α*, the next state will be* q_2*. For every* $q \in St, \alpha \in Act$*, we assume that either (1)* $\tau(q, \alpha, q') = 0$ *for all* q' *(i.e.,* α *is not enabled in* q*), or (2)* $\tau(q, \alpha, \cdot)$ *is a probability distribution.*

Additionally, we define $act(q) = \{\alpha \in Act \mid \exists q'.\tau(q, \alpha, q') > 0\}$ *as the set of enabled actions in* q*.*

A policy is a conditional plan that specifies future actions of the decision-making agent. Policies can be stochastic as well, thus allowing for randomness in the agent's play.

Definition 3 *A* policy *(or* strategy*) in a Markov decision process* $\mathcal{M} = \langle St, Act, \tau, \pi \rangle$ *is a function* $s : States \times Act \to [0, 1]$ *that assigns each state* q *with a probability distribution over the enabled actions* $act(q)$*. That is,* $s(q, \alpha) \in [0, 1]$ *for all* $q \in St, \alpha \in act(q)$*, and* $\sum_{\alpha \in act(q)} s(q, \alpha) = 1$*. Values of* $s(q, \alpha)$ *for* $\alpha \notin act(q)$ *are irrelevant.*

Policy s *is deterministic (or* pure*) iff for each state* q *it specifies a single action* α *(i.e.,* $s(q, \alpha) = 1$*, and* $s(q, \alpha') = 0$ *for all the other* α'*). By abuse of notation, we will sometimes write* $s(q) = \alpha$ *instead of* $s(q, \alpha) = 1$ *for pure policies.*

The set of all policies in M *is denoted by* Σ_M*. The set of all pure policies in* M *is denoted by* σ_M*.*

Note that, if the agent's policy is fixed, a Markov decision process reduces to a Markov chain.

Definition 4 *Policy* $s : States \times Act \to [0, 1]$ instantiates MDP $\mathcal{M} = \langle St, Act, \tau, \pi \rangle$ *to a Markov chain* $\mathcal{M} \dagger s = \langle St', \tau', \pi' \rangle$ *with* $St' = St$*,* $\pi' = \pi$*, and* $\tau'(q, q') = \sum_{\alpha \in act(q)} s(q, \alpha)\, \tau(q, \alpha, q')$*.*

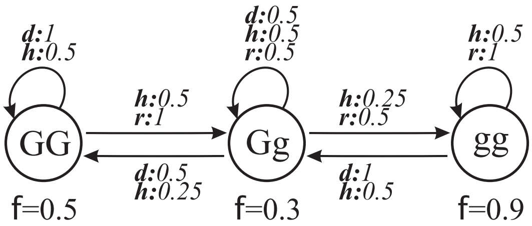

Figure 4: Markov decision process that allows for various mating policies

Example 2 (Gene model ctd.) *An extension of the "gene model" Markov chain from Example 1 is shown in Figure 4. Now, it is possible to mate the offspring with an animal that has dominant genes (action* d*), recessive genes (action* r*), or hybrid genes (action* h*). Note that the pure policy* $s(GG) = s(Gg) = s(gg) = h$ *instantiates the* MDP *to the Markov chain from Figure 3.*

Another example of instantiating Markov decision processes is shown in Figure 6.

Markov decision processes have been further extended to include imperfect information of the decision maker (POMDP: Partially Observable Markov Decision Processes [12]), and information about other agents acting in the same environment (MMDP: multi-agent Markov decision processes [8]). Extensions that handle both partial observability and multi-agent interplay were also considered [38].

4 MTL_0: A Logic of Markov Chains

In this section we present our first take on Markov Temporal Logic (MTL), a logic that allows for flexible reasoning about outcomes of agents acting in stochastic environments. The core of the logic is called MTL_0, and addresses outcomes of Markov chains. Intuitively, MTL_0 is a quantitative analogue of the branching-time logic CTL* [15, 14]; we will formalize (and prove) this claim later, in Section 4.4.

Operators of MTL_0 include path quantifiers $\mathsf{E}, \mathsf{A}, \mathsf{M}$ for the maximal, minimal, and average outcome of a set of temporal paths, respectively, and temporal operators $\Diamond, \Box, \mathsf{m}$ for the minimal, maximal, and average outcome along a given path.[3] Propositional operators follow the same pattern. Besides $\vee, \wedge$ for maximization and minimization of outcomes obtained from different utility channels or related to different goals, we use (after [10, 11]) the "weighted

[3] The temporal operators will allow to discount future outcomes with a discount factor c. Also, we will introduce the "until" operator $\mathcal{U}$, which is more general than $\Diamond$.

average" operator $\oplus$ which will prove useful when we formulate e.g. fixpoint properties of temporal operators with discount.

One of the novel contributions in our approach is the introduction of "defuzzification" operator $\preccurlyeq$. $\varphi_1 \preccurlyeq \varphi_2$ yields "true" if the outcome of φ_1 is less or equal to φ_2, and "false" otherwise. This provides a neat two-valued interface to the logic. Among other advantages, it allows to define the classical computational problems of validity, satisfiability and model checking for MTL.

Note that $\preccurlyeq$ can be seen as a kind of "crisp" material implication. One may also consider various forms of fuzzy implication (e.g., the one derived from $\varphi_1 \to \varphi_2 \equiv \neg\varphi_1 \vee \varphi_2$), but we are not sure if they are really useful (and if so, in which context).

4.1 Syntax of MTL$_0$

The syntax of MTL$_0$ (parameterized by a set of utility fluents Π) is defined as follows:

$$\begin{array}{rcl} \varphi & ::= & p \mid \neg\varphi \mid \varphi \wedge \varphi \mid \varphi \oplus_c \varphi \mid \varphi \preccurlyeq \varphi \mid \mathsf{E}\gamma \mid \mathsf{M}\gamma, \\ \gamma & ::= & \varphi \mid \neg\gamma \mid \gamma \wedge \gamma \mid \bigcirc_c \gamma \mid \Box_c\gamma \mid \gamma\,\mathcal{U}_c\,\gamma \mid \mathsf{m}_c\gamma. \end{array}$$

where $p \in \Pi$ is a utility fluent, and $c \in (0, 1]$ is a discount factor. We will use the symbol $\mathcal{L}state(\Pi)$ to denote the set of "state formulae" φ (i.e., the set of proper formulae of MTL$_0$), and $\mathcal{L}path(\Pi)$ to denote the set of "path formulae" γ.

Additionally, we define the Boolean constants T, F (standing for "true" and "false"), disjunction, and the "sometime" temporal operator $\Diamond$ as below. Except for T, all of them are just standard definitions that can be found in any textbook on temporal logic. We will show in Section 4.2 that their semantics corresponds to our intuition also in this setting.

- $\mathrm{T} \equiv p \preccurlyeq p$,
- $\mathrm{F} \equiv \neg\mathrm{T}$,
- $\varphi_1 \vee \varphi_2 \equiv \neg(\neg\varphi_1 \wedge \neg\varphi_2)$,
- $\mathsf{A}\gamma \equiv \neg\mathsf{E}\neg\gamma$,
- $\gamma_1 \vee \gamma_2 \equiv \neg(\neg\gamma_1 \wedge \neg\gamma_2)$,
- $\Diamond_c\gamma \equiv \mathrm{T}\,\mathcal{U}_c\,\gamma$,
- $\varphi_1 \cong \varphi_2 \equiv (\varphi_1 \preccurlyeq \varphi_2) \wedge (\varphi_2 \preccurlyeq \varphi_1)$.

We may also use the following shorthands for discount-free versions of temporal operators: $\bigcirc \equiv \bigcirc_1, \Diamond \equiv \Diamond_1, \Box \equiv \Box_1, \mathcal{U} \equiv \mathcal{U}_1$.

Example 3 *The following* MTL$_0$ *formulae define some interesting characteristics of the breeding process from Example 1:* $\mathsf{M}\mathsf{m}_{0.9}\mathsf{f}$ *(expected average fitness with time* *discount* 0.9*),* $\mathsf{Am}_{0.9}\mathsf{f}$ *(guaranteed average fitness with the same discount factor),* $\mathsf{M}\Box\mathsf{f}$ *(expected minimal future fitness), and* $\mathsf{A}\Diamond\mathsf{f}$ *(guaranteed maximal fitness).*

Note that the syntax of MTL$_0$ extends that of branching-time logic CTL*, where only operators A, E, $\bigcirc$, $\Box$, $\Diamond$, $\mathcal{U}$, and the Boolean connectives are used.

4.2 Semantics of MTL$_0$

The main idea behind MTL$_0$ is to treat formulae in a sufficiently general way, so that they can represent both quantitative utilities and qualitative truth values referring to something which is *completely* true or false (like a task that has been completely achieved). On a basic level, this made us define propositional valuations to yield values from $\hat{U}$. In consequence, a propositional letter p may refer to a distribution of rewards for a particular task among states of a model; it can also deem some states as ones in which the task is completely achieved or ones in which the task has been failed. Besides advantages in terms of modeling, this allows to freely mix qualitative and quantitative properties, which (hopefully) makes the resulting semantics elegant and powerful. Thus, we are going to treat complex formulae as fluents, just like the atomic utility fluents from Π, through a valuation function that assigns formulae with extended utility values from $\hat{U}$.

Let $M = \langle St, \tau, \pi\rangle$ be a Markov chain over domain $D = \langle U, \top, \bot, {}^{-}\rangle$ and a set of utility fluents Π. The truth value of formulae in M is determined by the valuation function $[\cdot] : (St \times \mathcal{L}state(\Pi)) \cup (\mathcal{R} \times \mathcal{L}path(\Pi)) \to \hat{U}$, defined below. We will omit M in $[\cdot]_{M,q}$, $[\cdot]_{M,\lambda}$ when the model is clear from the context.

- $[p]_q = \pi(p, q)$, for $p \in \Pi$;
- $[\neg\varphi]_q = \overline{[\varphi]_q}$;
- $[\varphi_1 \wedge \varphi_2]_q = \min([\varphi_1]_q, [\varphi_2]_q)$;
- $[\varphi_1 \oplus_c \varphi_2]_q = (1-c)\cdot[\varphi_1]_q + c\cdot[\varphi_2]_q$. Note that $\varphi_1 \oplus_c \varphi_2$ is *not* commutative, unless $c = 0.5$;
- $[\varphi_1 \preccurlyeq \varphi_2]_q = \top$ if $[\varphi_1]_q \leq [\varphi_2]_q$ and $\bot$ otherwise. That is, $\preccurlyeq$ "defuzzifies" formulae in the sense that it alsways yields a classical truth value $\top$ or $\bot$. Note also that $\preccurlyeq$ can be seen as a kind of material implication for multi-valued sentences;
- $[\varphi]_{M,\lambda} = [\varphi]_{M,\lambda[0]}$;
- $[\neg\gamma]_\lambda = \overline{[\gamma]_\lambda}$;
- $[\gamma_1 \wedge \gamma_2]_\lambda = \min([\gamma_1]_\lambda, [\gamma_2]_\lambda)$;
- $[\bigcirc_c \gamma]_\lambda = c \cdot [\gamma]_{\lambda[1..\infty]}$;
- $[\Box_c \gamma]_{M,\lambda} = \inf_{i=0,1,\dots}\{c^i[\gamma]_{M,\lambda[i..\infty]}\}$: the "always" operator minimizes utility along λ;

- $[\gamma_1 \mathcal{U}_c \gamma_2]_\lambda = \sup_{i=0,1,\dots} \{\min(\min_{0\le j<i}\{c^j[\gamma_1]_{\lambda[j..\infty]}\},\ c^i[\gamma_2]_{\lambda[i..\infty]})\}$;
- The Markovian temporal operator m_c produces the average discounted reward along the given run, which is in fact the cumulative discounted reward normalized with $1-c$:

$$[\mathsf{m}_c\gamma]_\lambda = \begin{cases} (1-c)\sum_{i=0}^{\infty} c^i[\gamma]_{\lambda[i\dots\infty]} & \text{if } c<1 \\ \lim_{i\to\infty}\frac{1}{i+1}\sum_{j=0}^{i}[\gamma]_{\lambda[i\dots\infty]} & \text{if } c=1 \end{cases}$$

- $[\mathsf{E}\gamma]_q = \sup_{\lambda\in\mathcal{R}(q)}\{[\gamma]_\lambda\}$: the existential path quantifier maximizes the utility across paths starting from q;
- The Markovian path quantifier $\mathsf{M}\gamma$ produces the expected truth value γ across all the possible runs (from now on). Given M, q, we first define the probability space $\langle\mathcal{R}(q), \mathcal{H}(q), pr\rangle$ induced by the next-state transition probabilities τ (cf. also [10, 24, 37]). In this space, elementary outcomes are runs from $\mathcal{R}(q)$, events are sets of runs that share the same finite prefix (i.e., ones from $\mathcal{H}(q)$), and the probability measure $pr : \mathcal{H}(q) \to [0,1]$ is defined as $pr(q_0 \dots q_1) = \tau(q_0, q_1)\cdot\dots\cdot\tau(q_{i-1}, q_i)$. Then, we use the valuation of γ as the random variable; the truth value of $\mathsf{M}\gamma$ is defined as its expected value:

$$\begin{aligned}[\mathsf{M}\gamma]_q &= E[\gamma] = \int_{\mathcal{R}(q)}[\gamma]_\lambda pr(\lambda)d\lambda = \lim_{k\to\infty}\sum_{h\in\mathcal{H}^k(q)}[\gamma]_{\lambda(h)}pr(\lambda(h)) \\ &= \lim_{k\to\infty}\sum_{h\in\mathcal{H}^k(q)}[\gamma]_{\lambda(h)}\tau(h[0],h[1])\cdot\dots\cdot\tau(h[k-1],h[k]).\end{aligned}$$

Example 4 *Below, we present the valuations of the* MTL$_0$ *formulae from Example 3 for the breeding process from Figure 3.* $[\mathsf{Mm}_{0.9}\mathsf{f}]_{GG} = 0.484$, $[\mathsf{Mm}_{0.9}\mathsf{f}]_{Gg} = 0.480$, *and* $[\mathsf{Mm}_{0.9}\mathsf{f}]_{gg} = 0.554$*; i.e., the expected average fitness with time discount* 0.9 *is* $0.484, 0.480, 0.554$ *if we start with dominant, hybrid, and recessive genes, respectively. Moreover,* $[\mathsf{Am}_{0.9}\mathsf{f}]_{GG} = 0.32$, $[\mathsf{Am}_{0.9}\mathsf{f}]_{Gg} = 0.3$, *and* $[\mathsf{Am}_{0.9}\mathsf{f}]_{gg} = 0.36$*: the guaranteed average fitness (with discount) is* $0.32, 0.3, 0.36$*, respectively. Finally, the expected minimal undiscounted fitness* $[\mathsf{M}\Box\mathsf{f}]_q = 0.3$ *for all states* q*, and the guaranteed maximal fitness* $[\mathsf{A}\Diamond\mathsf{f}]_q = 0.3$ *for all states* q.

Proposition 1 *We note that the derived operators have the following semantic characteristics:*

1. $[\mathrm{T}]_{M,q} = \top$ *for every* M, q;
2. $[\mathrm{F}]_{M,q} = \bot$ *for every* M, q;
3. $[\varphi_1 \vee \varphi_2]_{M,q} = \max([\varphi_1]_{M,q}, [\varphi_2]_{M,q})$;
4. $[\gamma_1 \vee \gamma_2]_{M,\lambda} = \max([\gamma_1]_{M,\lambda}, [\gamma_2]_{M,\lambda})$. *That is, disjunction is indeed a "truth value maximizer" for both state and path formulae;*

5. $[\mathsf{A}\gamma]_{M,q} = \inf_{\lambda\in\mathcal{R}(q)}\{[\gamma]_{M,\lambda}\}$: *the universal path quantifier minimizes the utility across paths starting from* q*;*
6. $[\Diamond_c\gamma]_{M,\lambda} = \sup_{i=0,1,\ldots}\{c^i[\gamma]_{M,\lambda[i..\infty]}\}$: *the "sometime" operator maximizes discounted utility along* λ*;*
7. $[\varphi_1 \cong \varphi_2]_{M,q} = \top$ *if* $[\varphi_1]_{M,q} = [\varphi_2]_{M,q}$*, and* $\bot$ *otherwise. That is,* $\cong$ *captures a very strong notion of equivalence between multi-valued sentences, namely that both sentences have exactly the same truth value.*

Proof. Let M be an arbitrary Markov chain.

1. $[\mathrm{T}]_q = [p \preccurlyeq p]_q = \top$.
2. $[\mathrm{F}]_q = \overline{\top} = \bot$.
3. $[\varphi_1 \vee \varphi_2]_q = [\neg(\neg\varphi_1 \wedge \neg\varphi_2)]_q = \overline{\min(\overline{[\varphi_1]_q}, \overline{[\varphi_2]_q})} = \max([\varphi_1]_q, [\varphi_2]_q)$ by the properties of quasi-boolean complement.
4. $[\gamma_1 \vee \gamma_2]_\lambda = [\neg(\neg\gamma_1 \wedge \neg\gamma_2)]_\lambda = \overline{\min(\overline{[\gamma_1]_\lambda}, \overline{[\gamma_2]_\lambda})} = \max([\gamma_1]_\lambda, [\gamma_2]_\lambda)$.
5. $[\mathsf{A}\gamma]_q = [\neg\mathsf{E}\neg\gamma]_q = \overline{\sup_{\lambda\in\mathcal{R}(q)}\{\overline{[\gamma]_\lambda}\}} = \inf_{\lambda\in\mathcal{R}(q)}\{[\gamma]_\lambda\}$.
6. $[\Diamond_c\gamma]_\lambda = [\mathrm{T}\,\mathcal{U}_c\,\gamma]_\lambda = \sup_{i=0,1,\ldots}\big\{\min(\,\min_{0\le j<i}\{c^j\top\},\ c^i[\gamma_2]_{\lambda[i..\infty]})\big\} = \sup_{i=0,1,\ldots}\{\min(c^{i-1}\top, c^i[\gamma_2]_{\lambda[i..\infty]})\} = \sup_{i=0,1,\ldots}\{c^i[\gamma]_{\lambda[i..\infty]}\}$.
7. $[\varphi_1 \cong \varphi_2]_{M,q} = \min([\varphi_1 \preccurlyeq \varphi_2]_{M,q}, [\varphi_2 \preccurlyeq \varphi_1]_{M,q}) = \top$ if $[\varphi_1]_{M,q} \le [\varphi_2]_{M,q}$ and $[\varphi_2]_{M,q} \le [\varphi_1]_{M,q}$, and $\bot$ otherwise.

■

The undiscounted versions of temporal operators "always" and "sometime" have the usual relationship, but it does not transfer to the discounted case. Moreover, discounted "always" is trivial for many domains.

Proposition 2

1. $[\Box\gamma]_{M,\lambda} = [\neg\Diamond\neg\gamma]_{M,\lambda}$*,*
2. $[\Box_c\gamma]_{M,\lambda} = 0$ *if* $c < 1$ *and* $\hat{U} \subseteq \mathbb{R}_+ \cup \{0\}$.

Proof. Let M be an arbitrary Markov chain, and λ a run in M.

1. $[\Box_c\gamma]_\lambda = \inf_{i=0,1,\ldots}\{c^i[\gamma]_{\lambda[i..\infty]}\}$, and $[\neg\Diamond_c\neg\gamma]_{M,\lambda} = \overline{\sup_{i=0,1,\ldots}\{c^i\overline{[\gamma]_{\lambda[i..\infty]}}\}} = \inf_{i=0,1,\ldots}\{\overline{c^i\overline{[\gamma]_{\lambda[i..\infty]}}}\}$. These two are equal if $\overline{c^i\overline{u}} = c^i u$, which is the case for $c = 1$, but in general it does not hold.
2. We recall the assumption that M has a finite number of states. Then, there is a finite number of utility values from $\hat{U}$ that actually occur in M (as values of π). We take the maximal one and denote it by u_{max}. Since $\hat{U} \subseteq \mathbb{R}_+ \cup \{0\}$, we have that $u_{max} \in \mathbb{R}_+ \cup \{0\}$. It is easy to see that,

for each state formula φ, the value of φ in any state q cannot be higher than u_{max}.

Now, $[\Box_c\gamma]_\lambda = \inf_{i=0,1,\dots}\{c^i[\gamma]_{\lambda[i..\infty]}\} \leq \inf_{i=0,1,\dots}\{c^i \sup_{\lambda'\in\mathcal{R}(\lambda[i])}[\gamma]_{\lambda'}\} = \inf_{i=0,1,\dots}\{c^i[\mathsf{E}\gamma]_{\lambda[i]}\} \leq \inf_{i=0,1,\dots}\{c^i u_{max}\} = 0$. But we have also that $[\Box_c\gamma]_\lambda \geq 0$, since all the utility values are non-negative. Thus, $[\Box_c\gamma]_\lambda = 0$.

■

4.3 Levels of Truth

Since every domain must include a distinguished value for the classical (complete) truth, validity of formulae can be defined in a straightforward way.

Definition 5 (Levels of validity) *Let M be a Markov chain, q a state in M, and φ a formula of* MTL$_0$. *Then:*

- *φ is* true in *M, q (written $M, q \models \varphi$) iff $[\varphi]_{M,q} = \top$.*
- *φ is* valid in *M (written $M \models \varphi$) iff it is true in every state of M.*
- *φ is* valid for Markov chains *(written $\models \varphi$) iff it is valid in every Markov chain M.*
- *Additionally, for path formulae γ, we can say that γ holds on run λ in Markov chain M (written $M, \lambda \models \gamma$) iff $[\gamma]_{M,\lambda} = \top$.*

Example 5 *Let M be the Markov chain from Figure 3 with additional utility fluents* $\mathsf{0.3}$, $\mathsf{0.32}$ *and* $\mathsf{0.36}$ *such that* $\pi(\mathsf{0.3}, q) = 0.3$, $\pi(\mathsf{0.32}, q) = 0.32$, *and* $\pi(\mathsf{0.36}, q) = 0.36$ *for all $q \in St$. Then, we have that* $M, GG \models \mathsf{Am}_{0.9}\mathsf{f} \cong \mathsf{0.32}$, $M, Gg \models \mathsf{Am}_{0.9}\mathsf{f} \cong \mathsf{0.3}$, *and* $M, gg \models \mathsf{Am}_{0.9}\mathsf{f} \cong \mathsf{0.36}$. *Moreover, the following formula is valid in M:* $M \models \mathsf{0.3} \preccurlyeq \mathsf{Am}_{0.9}\mathsf{f} \wedge \mathsf{Am}_{0.9}\mathsf{f} \preccurlyeq \mathsf{0.36}$.

Note that T is valid for Markov chains, while F is true in no M, q. Other examples of validities are: $\mathsf{A}\Box\gamma \cong \mathsf{A}\neg\Diamond\neg\gamma$, $\mathsf{E}\Box\gamma \cong \mathsf{E}\neg\Diamond\neg\gamma$ etc. (cf. Proposition 2.1).

Definition 5 enables the traditional view of MTL$_0$ that identifies "the logic" with the set of valid formulae of that logic. Moreover, it allows to define the typical decision problems for MTL$_0$ in a natural way:

- Given a formula φ, the *validity problem* asks if $\models \varphi$;
- Given a formula φ, the *satisfiability problem* asks if there are M, q such that $M, q \models \varphi$;
- Given a model M, state q and formula φ, the *model checking problem* asks if $M, q \models \varphi$ (this variant is sometimes called *local* model checking). Alternatively, one can use the definition adopted in [10, 11], where the

output of model checking is the truth value $[\varphi]_{M,q}$. For *global model checking*, the input consists of M and φ, and we ask for the exact set of states such that $M, q \models \varphi$.

An important corollary of Proposition 1.7 is that the notion of equivalence defined by $\cong$ is strong enough to make equivalent (sub)formulae completely interchangeable on all levels of validity.

Corollary 3 *If $M, q \models \varphi_1 \cong \varphi_2$, and ψ' is obtained from ψ through replacing an occurrence of φ_1 by φ_2, then $M, q \models \psi$ iff $M, q \models \psi'$.*

4.4 Transition Systems as Markov Chains. Correspondence between MTL$_0$ and CTL*

Markov chains can be seen as generalizations of transition systems, where quantitative information is added via non-classical values of atomic statements and probabilities of transitions. As action labels are absent in Markov chains, these in fact generalize *unlabeled* transition systems (UTS). In this section, we redefine UTS as a proper subclass of Markov chains, in which all the fluents can accept only classical truth values.

Definition 6 *Let M be a Markov chain. Formula φ is* propositional *in M iff it can take only the values of $\top, \bot$, i.e., $[\varphi]_{M,q} \in \{\top, \bot\}$ for all $q \in St$.*

Propositions have a simple characterization for Markov chains.

Proposition 4 *Let M be a Markov chain and φ a formula of* MTL$_0$. *Then φ is propositional in M iff formula $(\varphi \cong \mathrm{F}) \vee (\varphi \cong \mathrm{T})$ is valid in M.*

Proof. Straightforward. ∎

An *unlabeled transition system* can be defined as a Markov chain with only propositional fluents. We also require that the domain of the MC has more than one element to make sure that the classical truth values $\top, \bot$ are distinct in the model.[4] This way, we obtain the class of models that are used for qualitative branching-time logics, i.e. CTL and CTL*. Of course, when interpreting formulae of CTL*, one must also ignore the probabilities that are present in Markov chains. The next two propositions show that MTL$_0$ strictly generalizes CTL*.

Proposition 5 *Let M be a transition system, q a state in M, and φ a (state) formula of* CTL*. *Then, $M, q \models_{\text{MTL}_0} \varphi$ iff $M, q \models_{\text{CTL*}} \varphi$.*

Likewise, for every transition system M, path λ in M, and path formula γ of CTL*, *we have that $M, \lambda \models_{\text{MTL}_0} \gamma$ iff $M, \lambda \models_{\text{CTL*}} \gamma$.*

[4] Note that the requirement is important, as e.g. no CTL* formula $\neg\varphi$ will hold if $\top = \bot$.

The following lemma states that CTL* formulae take only classical truth values in transition systems. We will need it to prove the case for negation.

Lemma 6 *For every state formula φ of CTL*, every transition system M, and every state q in M, it holds that $[\varphi]_{M,q} \in \{\top, \bot\}$. Likewise, for every path formula γ of CTL*, every transition system M, and every path λ in M, we have that $[\gamma]_{M,\lambda} \in \{\top, \bot\}$.*

Proof. Straightforward induction on the structure of φ/γ. ∎

Proof of Proposition 5. Structural induction on the structure of φ/γ.

Case $\varphi \equiv p$: $M, q \models_{\text{MTL}_0} p$ iff $[p]_{M,q} = \top$ iff $\pi(p, q) = \top$ iff $M, q \models_{\text{CTL*}} p$.

Case $\varphi \equiv \neg\varphi'$: $M, q \models_{\text{MTL}_0} \neg\varphi'$ iff $[\neg\varphi']_{M,q} = \top$ iff $\overline{[\varphi']_{M,q}} = \top$ iff $\overline{\overline{[\varphi']_{M,q}}} = \overline{\top}$ iff $[\varphi']_{M,q} = \bot$[5] iff $[\varphi']_{M,q} \neq \top$[6] iff $M, q \not\models_{\text{MTL}_0} \varphi'$ iff (by induction) $M, q \not\models_{\text{CTL*}} \varphi'$ iff $M, q \models_{\text{CTL*}} \neg\varphi'$.

Case $\varphi \equiv \varphi_1 \wedge \varphi_2$: $M, q \models_{\text{MTL}_0} \varphi_1 \wedge \varphi_2$ iff $[\varphi_1 \wedge \varphi_2]_{M,q} = \top$ iff $[\varphi_1]_{M,q} = \top$ and $[\varphi_2]_{M,q} = \top$ iff $M, q \models_{\text{MTL}_0} \varphi_1$ and $M, q \models_{\text{MTL}_0} \varphi_2$ iff (by induction) $M, q \models_{\text{CTL*}} \varphi_1$ and $M, q \models_{\text{CTL*}} \varphi_2$ iff $M, q \models_{\text{CTL*}} \varphi_1 \wedge \varphi_2$.

Case $\varphi \equiv \mathsf{E}\gamma$: $M, q \models_{\text{MTL}_0} \mathsf{E}\gamma$ iff $[\mathsf{E}\gamma]_{M,q} = \top$ iff $\sup_{\lambda \in \mathcal{R}(q)}\{[\gamma]_{M,\lambda}\} = \top$ iff $\exists_{\lambda \in \mathcal{R}(q)} [\gamma]_{M,\lambda} = \top$ iff $\exists_{\lambda \in \mathcal{R}(q)} M, \lambda \models_{\text{MTL}_0} \gamma$ iff (by induction) $\exists_{\lambda \in \mathcal{R}(q)} M, \lambda \models_{\text{CTL*}} \gamma$ iff $M, q \models_{\text{CTL*}} \mathsf{E}\gamma$.

Cases $\gamma \equiv \neg\gamma'$, $\gamma \equiv \gamma_1 \wedge \gamma_2$: similar to the analogous state formulae.

Case $\gamma \equiv \bigcirc\gamma'$: $M, \lambda \models_{\text{MTL}_0} \bigcirc\gamma'$ iff $[\bigcirc\gamma']_{M,\lambda} = \top$ iff $[\gamma']_{M,\lambda[1..\infty]} = \top$ iff $M, \lambda[1..\infty] \models_{\text{MTL}_0} \gamma'$ iff (by induction) $M, \lambda[1..\infty] \models_{\text{CTL*}} \gamma'$ iff $M, \lambda \models_{\text{CTL*}} \bigcirc\gamma'$.

Cases $\gamma \equiv \Box\gamma'$, $\gamma \equiv \gamma_1 \,\mathcal{U}\, \gamma_2$: analogous.

∎

Proposition 7 *There is a transition system M with states q, q' which cannot be distinguished by any CTL* formula, and can be distinguished by a formula of MTL$_0$.*

Proof. Consider the transition system in Figure 5. Note that states q_1, q_2 are bisimilar under CTL* bisimulation, so the same CTL* properties hold in both states (cf. e.g. [30]). On the other hand, we have that $[\mathsf{Mmp}]_{q_1} = 0.5 = [\mathsf{Em}_{0.5}\mathsf{p}]_{q_1}$, and $[\mathsf{Mmp}]_{q_2} = 0.1 \neq 0.5 = [\mathsf{Em}_{0.5}\mathsf{p}]_{q_2}$. Thus, for $\varphi \equiv \mathsf{Mmp} \cong \mathsf{Em}_{0.5}\mathsf{p}$, we have $q_1 \models \varphi$ and $q_2 \not\models \varphi$ (and even $q_2 \models \neg\varphi$). ∎

The above example shows – unsurprisingly – that a proper notion of bisimulation for Markov chains must take into account transition probabilities.

[5] $\overline{[\varphi']_{M,q}} = \top$ implies $\overline{\overline{[\varphi']_{M,q}}} = \overline{\top}$ implies $[\varphi']_{M,q} = \bot$ implies $\overline{[\varphi']_{M,q}} = \overline{\bot}$ implies $\overline{[\varphi']_{M,q}} = \top$.

[6] Left to right: by the requirement that $\top, \bot$ are distinct; right to left: by Lemma 6.

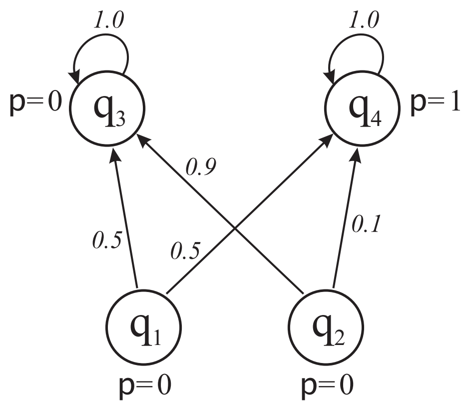

Figure 5: MTL_0 vs. CTL*: probabilities matter!

Remark 8 *Note that, unlike in two-valued logic, $M, q \not\models \varphi$ does not necessarily imply that $M, q \models \neg\varphi$. The first requires only that $[\varphi]_{M,q} \neq \top$, while the latter means that $[\varphi]_{M,q} = \bot$.*

4.5 State-Based MTL_0

"CTL without star" (or "vanilla" CTL [9]) is the most often used variant of computation tree logic, mainly due to the complexity of its model checking problem (linear with respect to the number of transitions in the model and the length of the formula). The fact that its semantics can be defined entirely in relation to states (rather than both states and runs) also plays a role. "Vanilla" CTL can be seen as a syntactic restriction of CTL*, in which every temporal modality is preceded by exactly one path quantifier. In this section, we consider a similar syntactic restriction on MTL_0; we call it state-based MTL_0.

Definition 7 State-based MTL_0 *(sMTL_0 in short) is given by the following grammar (where $p \in \Pi$ stands for utility fluents, and $c \in (0,1]$ for discount factors):*

$$\begin{aligned} \varphi &::= p \mid \neg\varphi \mid \varphi \wedge \varphi \mid \varphi \oplus_c \varphi \mid \varphi \preccurlyeq \varphi \mid \mathsf{E}\gamma \mid \mathsf{A}\gamma \mid \mathsf{M}\gamma, \\ \gamma &::= \bigcirc_c \varphi \mid \Box_c \varphi \mid \varphi\, \mathcal{U}_c\, \varphi \mid \mathsf{m}_c \varphi. \end{aligned}$$

So, the structure of a typical sMTL_0 formula is as follows: first a path quantifier, then a temporal quantifier, and then a propositional formula or a nested sMTL_0 formula. As the truth value of the nested formula depends only of the state of evaluation, it is easy to define the semantics of sMTL_0 entirely in terms of states, just like for "vanilla" CTL (giving one clause for each combination of a path and temporal modality). We leave this as an exercise for the interested reader.

Lemma 9 shows that $\mathsf{E}\bigcirc_c \varphi$, $\mathsf{A}\bigcirc_c \varphi$, and $\mathsf{M}\bigcirc_c \varphi$ implement the discounted maximal, minimal, and expected value of φ in the next moment, respectively. Proposition 10 presents fixpoint characterizations for most modalities of sMTL$_0$. The results from [10, 11] suggest that $\mathsf{M}\Box_c$ and $\mathsf{M}\mathcal{U}_c$ do not have fixpoint characterizations, but this remains to be formally proven.

Lemma 9 *Let φ be a formula of* sMTL$_0$*. Then:*

1. $[\mathsf{E}\bigcirc_c \varphi]_q = c \max_{q' \in \tau(q)} [\varphi]_{q'}$;
2. $[\mathsf{A}\bigcirc_c \varphi]_q = c \min_{q' \in \tau(q)} [\varphi]_{q'}$;
3. $[\mathsf{M}\bigcirc_c \varphi]_q = c \sum_{q' \in \tau(q)} \tau(q, q')[\varphi]_{q'}$.

Proof.

1. $[\mathsf{E}\bigcirc_c \varphi]_q = \sup_{\lambda \in \mathcal{R}(q)}\{c[\varphi]_{\lambda[1]}\} = c \sup_{q' \in \tau(q)}\{[\varphi]_{q'}\}$.
2. Analogous.
3. $[\mathsf{M}\bigcirc_c \varphi]_q = \lim_{k\to\infty} \sum_{h \in \mathcal{H}^k(q)} [\bigcirc_c \varphi]_{\lambda(h)} pr(h) =$ $\lim_{k\to\infty} \sum_{h \in \mathcal{H}^k(q)} c[\varphi]_{h[1]} pr(h) =$ $c \lim_{k\to\infty} \sum_{q' \in \tau(q)} \sum_{h' \in \mathcal{H}^{k-1}(q')} [\varphi]_{h[1]} \tau(h[0], h[1]) \cdot \ldots \cdot \tau(h[k-1], h[k]) =$ $c \lim_{k\to\infty} \sum_{q' \in \tau(q)} [\varphi]_{q'} \tau(q, q') \sum_{h' \in \mathcal{H}^{k-1}(q')} pr(h') = c \sum_{q' \in \tau(q)} [\varphi]_{q'} \tau(q, q')$.

■

Proposition 10 *The following formulae of* sMTL$_0$ *are valid:*

1. $\mathsf{E}\Box_c\varphi \cong \varphi \wedge \mathsf{E}\bigcirc_c \mathsf{E}\Box_c\varphi$;
2. $\mathsf{A}\Box_c\varphi \cong \varphi \wedge \mathsf{A}\bigcirc_c \mathsf{A}\Box_c\varphi$;
3. $\mathsf{E}\varphi_1 \mathcal{U}_c \varphi_2 \cong \varphi_2 \vee \varphi_1 \wedge \mathsf{E}\bigcirc_c \mathsf{E}\varphi_1 \mathcal{U}_c \varphi_2$;
4. $\mathsf{A}\varphi_1 \mathcal{U}_c \varphi_2 \cong \varphi_2 \vee \varphi_1 \wedge \mathsf{A}\bigcirc_c \mathsf{A}\varphi_1 \mathcal{U}_c \varphi_2$;
5. $\mathsf{Em}_c\varphi \cong \varphi \oplus_c \mathsf{E}\bigcirc \mathsf{Em}_c\varphi$;
6. $\mathsf{Am}_c\varphi \cong \varphi \oplus_c \mathsf{A}\bigcirc \mathsf{Am}_c\varphi$;
7. $\mathsf{Mm}_c\varphi \cong \varphi \oplus_c \mathsf{M}\bigcirc \mathsf{Mm}_c\varphi$.

In order to prove Proposition 10, we will need the following technical lemma.

Lemma 11 *Let c be a constant and x a variable. Then:*

1. $\sup_x\{\min(c, f(x))\} = \min(c, \sup_x\{f(x)\})$.
2. $\inf_x\{\max(c, f(x))\} = \max(c, \inf_x\{f(x)\})$.

Proof.

1. $\sup_x\{\min(c, f(x))\} = \sup_x \left\{ \begin{array}{ll} c & \text{if } c \le f(x) \\ f(x) & \text{else} \end{array} \right\}$
   $= \left\{ \begin{array}{ll} c & \text{if } c \le \sup_x\{f(x)\} \\ \sup_x\{f(x)\} & \text{else} \end{array} \right\} = \min(c, \sup_x\{f(x)\}).$
2. Analogous.

■

Proof of Proposition 10.

1. $[\mathsf{E}\Box_c\varphi]_q = \sup_{\lambda\in\mathcal{R}(q)}\{\inf_{i=0,1,\ldots}\{c^i[\varphi]_{\lambda[i]}\}\} =$
   $\sup_{\lambda\in\mathcal{R}(q)}\{\min(c^0[\varphi]_{\lambda[0]}, \inf_{i=1,2,\ldots}\{c^i[\varphi]_{\lambda[i]}\})\} =$
   $\sup_{\lambda\in\mathcal{R}(q)}\{\min([\varphi]_q, c\inf_{i=0,1,\ldots}\{c^i[\varphi]_{\lambda[i-1]}\})\} =$
   $\sup_{q'\in\tau(q)}\sup_{\lambda'\in\mathcal{R}(q')}\{\min([\varphi]_q, c\inf_{i=0,1,\ldots}\{c^i[\varphi]_{\lambda'[i]}\})\} =$
   $\sup_{q'\in\tau(q)}\{\min([\varphi]_q, c\sup_{\lambda'\in\mathcal{R}(q')}\inf_{i=0,1,\ldots}\{c^i[\varphi]_{\lambda'[i]}\})\} =$
   $\min([\varphi]_q, \sup_{q'\in\tau(q)}\{c[\mathsf{E}\Box_c\varphi]_{q'}\}) = \min([\varphi]_q, [\mathsf{E}\bigcirc_c\mathsf{E}\Box_c\varphi]_q) =$
   $[\varphi\wedge\mathsf{E}\bigcirc_c\mathsf{E}\Box_c\varphi]_q.$
2. Analogous.
3. $[\mathsf{E}\varphi_1\,\mathcal{U}_c\,\varphi_2]_q = \sup_{\lambda\in\mathcal{R}(q)}\{\sup_{i=0,1,\ldots}\{\min(\min_{0\le j<i}\{c^j[\varphi_1]_{\lambda[j]}\}, c^i[\varphi_2]_{\lambda[i]})\}\} =$
   $\sup_{\lambda\in\mathcal{R}(q)}\{\max(\min(\min_{0\le j<0}\{c^j[\varphi_1]_{\lambda[j]}\}, c^0[\varphi_2]_{\lambda[0]}),$
   $\quad\sup_{i=1,2,\ldots}\{\min(\min_{0\le j<i}\{c^j[\varphi_1]_{\lambda[j]}\}, c^i[\varphi_2]_{\lambda[i]})\})\} =$
   $\max([\varphi_2]_q, \sup_{\lambda\in\mathcal{R}(q)}\{\sup_{i=1,2,\ldots}\{\min(\min_{0\le j<i}\{c^j[\varphi_1]_{\lambda[j]}\}, c^i[\varphi_2]_{\lambda[i]})\}\}) =$
   $\max([\varphi_2]_q, \sup_{\lambda\in\mathcal{R}(q)}\{\sup_{i=0,1,\ldots}\{\min(\min_{0\le j<i+1}\{c^j[\varphi_1]_{\lambda[j]}\}, c^{i+1}[\varphi_2]_{\lambda[i+1]})\}\}) =$
   $\max([\varphi_2]_q, \sup_{\lambda\in\mathcal{R}(q)}$
   $\quad\{\sup_{i=0,1,\ldots}\{\min(c^0[\varphi_1]_{\lambda[0]}, \min_{1\le j<i+1}\{c^j[\varphi_1]_{\lambda[j]}\}, c^{i+1}[\varphi_2]_{\lambda[i+1]})\}\}) =$
   $\max([\varphi_2]_q, \sup_{q'\in\tau(q)}\sup_{\lambda'\in\mathcal{R}(q')}$
   $\quad\{\sup_{i=0,1,\ldots}\{\min([\varphi_1]_q, \min_{1\le j<i+1}\{c^j[\varphi_1]_{\lambda'[j-1]}\}, c^{i+1}[\varphi_2]_{\lambda'[i]})\}\}) =$
   $\max([\varphi_2]_q, \min([\varphi_1]_q, \sup_{q'\in\tau(q)}\sup_{\lambda'\in\mathcal{R}(q')}$
   $\quad\{\sup_{i=0,1,\ldots}\{\min(\min_{0\le j<i}\{c^{j+1}[\varphi_1]_{\lambda'[j]}\}, c^{i+1}[\varphi_2]_{\lambda'[i]})\}\})) =$
   $\max([\varphi_2]_q, \min([\varphi_1]_q, c\sup_{q'\in\tau(q)}\sup_{\lambda'\in\mathcal{R}(q')}$
   $\quad\{\sup_{i=0,1,\ldots}\{\min(\min_{0\le j<i}\{c^j[\varphi_1]_{\lambda'[j]}\}, c^i[\varphi_2]_{\lambda'[i]})\}\})) =$
   $\max([\varphi_2]_q, \min([\varphi_1]_q, c\sup_{q'\in\tau(q)}\{[\mathsf{E}\varphi_1\,\mathcal{U}_c\,\varphi_2]_{q'}\})) =$
   $\max([\varphi_2]_q, \min([\varphi_1]_q, [\mathsf{E}\bigcirc_c\mathsf{E}\varphi_1\,\mathcal{U}_c\,\varphi_2]_q)) =$
   $[\varphi_2\vee\varphi_1\wedge\mathsf{E}\bigcirc_c\mathsf{E}\varphi_1\,\mathcal{U}_c\,\varphi_2]_q.$
4. Analogous.
5. Analogous to the proof of point 6.
6. **(a) Case** $c < 1$**:**
   $[\mathsf{Am}_c\varphi]_q = \inf_{\lambda\in\mathcal{R}(q)}\{(1-c)\sum_{i=0}^{\infty}c^i[\varphi]_{\lambda[i]}\} =$
   $\inf_{\lambda\in\mathcal{R}(q)}\{(1-c)c^0[\varphi]_{\lambda[0]} + (1-c)\sum_{i=1}^{\infty}c^i[\varphi]_{\lambda[i]}\} =$
   $(1-c)[\varphi]_q + \inf_{q'\in\tau(q)}\inf_{\lambda'\in\mathcal{R}(q')}\{(1-c)\sum_{i=1}^{\infty}c^i[\varphi]_{\lambda'[i-1]}\} =$
   $(1-c)[\varphi]_q + c\inf_{q'\in\tau(q)}\inf_{\lambda'\in\mathcal{R}(q')}\{(1-c)\sum_{i=0}^{\infty}c^i[\varphi]_{\lambda'[i]}\} =$

$(1-c)[\varphi]_q + c\inf_{q'\in\tau(q)}[\mathsf{Am}_c\varphi]_{q'} = (1-c)[\varphi]_q + c[\mathsf{A}\bigcirc\,\mathsf{Am}_c\varphi]_q =$
$[\varphi \oplus_c \mathsf{A}\bigcirc\,\mathsf{Am}_c\varphi]_q$;

(b) Case $c = 1$**:**
$[\mathsf{Am}\varphi]_q = \inf_{\lambda\in\mathcal{R}(q)}\{\lim_{i\to\infty}\frac{1}{i+1}\sum_{j=0}^{i}[\varphi]_{\lambda[j]}\} =$
$\inf_{q'\in\tau(q)}\inf_{\lambda'\in\mathcal{R}(q')}\{\lim_{i\to\infty}(\frac{[\varphi]_q}{i+1} + \frac{1}{i+1}\sum_{j=1}^{i}[\varphi]_{\lambda'[j-1]})\} =$
$\inf_{q'\in\tau(q)}\inf_{\lambda'\in\mathcal{R}(q')}\{\lim_{i\to\infty}(\frac{1}{i+1}\sum_{j=0}^{i-1}[\varphi]_{\lambda'[j]})\} =$
$\inf_{q'\in\tau(q)}\inf_{\lambda'\in\mathcal{R}(q')}\{\lim_{i\to\infty}(\frac{i+1}{i+2}\frac{1}{i+1}\sum_{j=0}^{i}[\varphi]_{\lambda'[j]})\} =$
$\inf_{q'\in\tau(q)}\{[\mathsf{Am}\varphi]_{q'}\} = [\mathsf{A}\bigcirc\,\mathsf{Am}\varphi]_q$.

7. **(a) Case** $c < 1$**:**
$[\mathsf{Mm}_c\varphi]_q = \lim_{k\to\infty}\sum_{h\in\mathcal{H}^k(q)}[\mathsf{m}_c\varphi]_{\lambda(h)}pr(h) =$
$\lim_{k\to\infty}\sum_{h\in\mathcal{H}^k(q)}(1-c)\big(\sum_{i=0}^{\infty}c^i[\varphi]_{\lambda(h)[i]}\big)pr(h) =$
$\lim_{k\to\infty}\sum_{h\in\mathcal{H}^k(q)}(1-c)\big(\sum_{i=0}^{k}c^i[\varphi]_{h[i]} + \sum_{i=k+1}^{\infty}c^i[\varphi]_{\lambda(h)[i]}\big)pr(h) =$
$\lim_{k\to\infty}\sum_{q'\in\tau(q)}\sum_{h'\in\mathcal{H}^{k-1}(q')}(1-c)$
$\quad\big(c^0[\varphi]_q + c\sum_{i=0}^{k-1}c^i[\varphi]_{h'[i]} + c\sum_{i=k}^{\infty}c^i[\varphi]_{\lambda(h')[i]}\big)\tau(q,q')pr(h') =$
$\lim_{k\to\infty}\sum_{q'\in\tau(q)}\tau(q,q')\big(\sum_{h'\in\mathcal{H}^{k-1}(q')}(1-c)[\varphi]_q pr(h') +$
$\quad c(1-c)(\sum_{h'\in\mathcal{H}^{k-1}(q')}\sum_{i=0}^{k-1}c^i[\varphi]_{h'[i]} + \sum_{i=k}^{\infty}c^i[\varphi]_{\lambda(h')[i]})pr(h')\big) =$
$\lim_{k\to\infty}\sum_{q'\in\tau(q)}\tau(q,q')\big((1-c)[\varphi]_q +$
$\quad c\sum_{h'\in\mathcal{H}^{k-1}(q')}(1-c)\sum_{i=0}^{\infty}c^i[\varphi]_{\lambda(h')[i]}pr(h')\big) =$
$\sum_{q'\in\tau(q)}\tau(q,q')\big((1-c)[\varphi]_q + c[\mathsf{Mm}_c\varphi]_{q'}\big) =$
$(1-c)[\varphi]_q\sum_{q'\in\tau(q)}\tau(q,q') + c\sum_{q'\in\tau(q)}\tau(q,q')[\mathsf{Mm}_c\varphi]_{q'} =$
$(1-c)[\varphi]_q + c[\mathsf{M}\bigcirc_1\,\mathsf{Mm}_c\varphi]_q = [\varphi\oplus_c\mathsf{M}\bigcirc\,\mathsf{Mm}_c\varphi]_q$;

(b) Case $c = 1$**:**
$[\mathsf{Mm}\varphi]_q = \lim_{k\to\infty}\sum_{h\in\mathcal{H}^k(q)}[\mathsf{m}\varphi]_{\lambda(h)}pr(h) =$
$\lim_{k\to\infty}\sum_{h\in\mathcal{H}^k(q)}\lim_{i\to\infty}\frac{1}{i+1}\sum_{j=0}^{i}[\varphi]_{\lambda(h)[j]}pr(h) =$
$\lim_{k\to\infty}\sum_{q'\in\tau(q)}\sum_{h'\in\mathcal{H}^{k-1}(q)}$
$\quad\lim_{i\to\infty}(\frac{[\varphi]_q}{i+1} + \tau(q,q')\frac{1}{i+1}\sum_{j=1}^{i}[\varphi]_{\lambda(h')[j-1]}pr(h')) =$
$\lim_{k\to\infty}\sum_{q'\in\tau(q)}\tau(q,q')\sum_{h'\in\mathcal{H}^{k-1}(q)}\lim_{i\to\infty}\frac{i+1}{i+2}\frac{1}{i+1}\sum_{j=0}^{i}[\varphi]_{\lambda(h')[j]}pr(h') =$
$\sum_{q'\in\tau(q)}\tau(q,q')[\mathsf{Mm}\varphi]_{q'} = [\mathsf{M}\bigcirc\,\mathsf{Mm}\varphi]_q$.

■

Example 6 *The characterizations enable computing the truth values of most* sMTL$_0$ *formulae by solving sets of simple equations. For instance, the valuations of formula* $\mathsf{Am}_{0.9}\mathsf{f}$ *for states* GG, Gg, gg *of the "gene model" Markov chain can be derived from the following equations:*

$$\begin{cases} [\mathsf{Am}_{0.9}\mathsf{f}]_{GG} &=& 0.1\cdot 0.5 + 0.9\min([\mathsf{Am}_{0.9}\mathsf{f}]_{GG}, [\mathsf{Am}_{0.9}\mathsf{f}]_{Gg}), \\ [\mathsf{Am}_{0.9}\mathsf{f}]_{Gg} &=& 0.1\cdot 0.3 + 0.9\min([\mathsf{Am}_{0.9}\mathsf{f}]_{GG}, [\mathsf{Am}_{0.9}\mathsf{f}]_{Gg}, [\mathsf{Am}_{0.9}\mathsf{f}]_{gg}), \\ [\mathsf{Am}_{0.9}\mathsf{f}]_{gg} &=& 0.1\cdot 0.9 + 0.9\min([\mathsf{Am}_{0.9}\mathsf{f}]_{Gg}, [\mathsf{Am}_{0.9}\mathsf{f}]_{gg}). \end{cases}$$

5 MTL$_1$: A Logic of Markov Decision Processes

The main aim of this paper is to offer a systematic study of temporal operators for Markov chains; the study was presented in the previous section. This section briefly shows how MTL$_0$ can be extended to strategic reasoning about Markov decision processes. We propose to use an explicit strategic quantifier $\langle\!\langle a \rangle\!\rangle$, similar to the *cooperation modality* from alternating-time temporal logic ATL [1, 2]. The intuitive meaning of $\langle\!\langle a \rangle\!\rangle \varphi$ is "the most that the decision maker can make out of φ". Note that there is always only one agent behind an MDP, so putting his name (e.g., "a") inside the operator is superfluous – but it will make the framework easier to extend to the multi-agent case in the future.

5.1 Syntax and Semantics of MTL$_1$

The syntax of MTL$_1$ is given by the following grammar:

$$\begin{array}{lcl} \vartheta & ::= & p \mid \neg\vartheta \mid \vartheta \wedge \vartheta \mid \vartheta \oplus_c \vartheta \mid \vartheta \preccurlyeq \vartheta \mid \langle\!\langle a \rangle\!\rangle \varphi, \\ \varphi & ::= & \neg\varphi \mid \varphi \wedge \varphi \mid \varphi \oplus_c \varphi \mid \mathsf{E}\gamma \mid \mathsf{M}\gamma, \\ \gamma & ::= & \vartheta \mid \neg\gamma \mid \gamma \wedge \gamma \mid \bigcirc_c \gamma \mid \Box_c \gamma \mid \gamma\, \mathcal{U}_c\, \gamma \mid \mathsf{m}_c \gamma. \end{array}$$

Note that a is just a fixed symbol and not a parameter of the strategic operator.

Let $\mathcal{M} = \langle St, Act, \tau, \pi \rangle$ be a Markov decision process over domain $D = \langle U, \top, \bot, \bar{\ } \rangle$ and a set of utility fluents Π. The truth value of formulae in M is determined by the valuation function $[\cdot]$ that extends the valuation of MTL$_0$ formulae from Section 4.2 as follows:

- $[p]_{\mathcal{M},q} = \pi(p, q)$, for $p \in \Pi$;
- $[\neg\vartheta]_{\mathcal{M},q} = \overline{[\vartheta]_{\mathcal{M},q}}$;
- $[\vartheta_1 \wedge \vartheta_2]_{\mathcal{M},q} = \min([\vartheta_1]_q, [\vartheta_2]_{\mathcal{M},q})$;
- $[\vartheta_1 \oplus_c \vartheta_2]_{\mathcal{M},q} = (1 - c) \cdot [\vartheta_1]_{\mathcal{M},q} + c \cdot [\vartheta_2]_{\mathcal{M},q}$;
- $[\vartheta_1 \preccurlyeq \vartheta_2]_{\mathcal{M},q} = \top$ if $[\vartheta_1]_{\mathcal{M},q} \leq [\vartheta_2]_{\mathcal{M},q}$ and $\bot$ otherwise;
- $[\langle\!\langle a \rangle\!\rangle \varphi]_{\mathcal{M},q} = \sup_{s \in \Sigma_M} \{[\varphi]_{M\dagger s,q}\}$;
- $[\vartheta]_{\mathcal{M}\dagger s,\lambda} = [\vartheta]_{\mathcal{M},\lambda[0]}$.

We use the same definitions of derived Boolean and temporal operators as in Section 4.1. Additionally, we define $\vartheta_1 \cong \vartheta_2 \equiv \vartheta_1 \preccurlyeq \vartheta_2 \wedge \vartheta_2 \preccurlyeq \vartheta_1$, and $[\![a]\!]\varphi \equiv \neg\langle\!\langle a \rangle\!\rangle\neg\varphi$. The following proposition shows that $[\![a]\!]\varphi$ implements the outcome of the worst possible policy with respect to φ.

Proposition 12 $[[\![a]\!]\varphi]_{\mathcal{M},q} = \inf_{s \in \Sigma_\mathcal{M}} \{[\varphi]_{\mathcal{M}\dagger s,q}\}$.

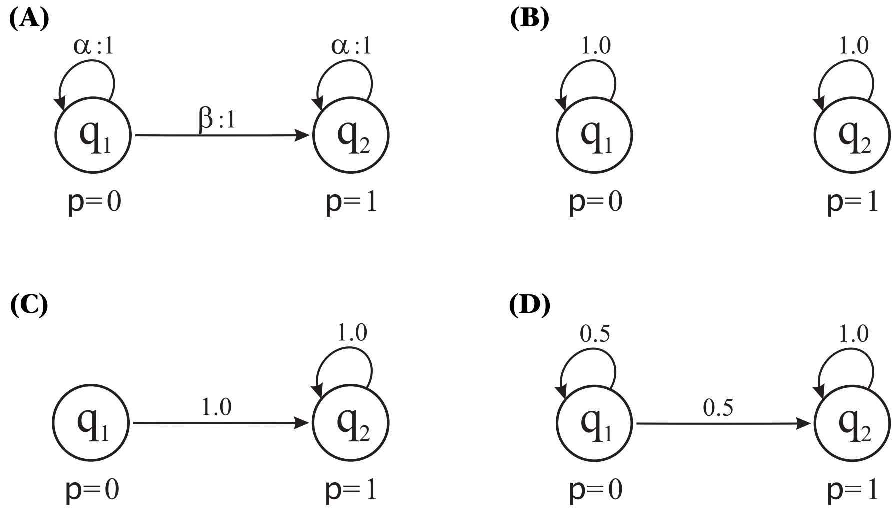

Figure 6: (A) A simple deterministic labeled transition system $\mathcal{M}$; (B) Instantiation of $\mathcal{M}$ by pure strategy $s_1(q_1) = \alpha$; (C) Instantiation of $\mathcal{M}$ by pure strategy $s_2(q_1) = \beta$; (D) Instantiation of $\mathcal{M}$ by strategy $s_3(q_1, \alpha) = s_3(q_1, \beta) = 0.5$.

Proof. $[[\![a]\!]\varphi]_{\mathcal{M},q} = \overline{[\langle\!\langle a\rangle\!\rangle \neg\varphi]_{\mathcal{M},q}} = \overline{\sup_{s\in\Sigma_\mathcal{M}}\{\overline{[\varphi]_{\mathcal{M}\dagger s,q}}\}} = \inf_{s\in\Sigma_\mathcal{M}}\{[\varphi]_{\mathcal{M}\dagger s,q}\}$. ■

Example 7 *Let $\mathcal{M}$ be the "gene model"* MDP *from Figure 4. Then, we have e.g.* $[\langle\!\langle a\rangle\!\rangle \mathsf{Mm}_{0.9}\mathsf{f}]_{GG} = 0.762$, $[\langle\!\langle a\rangle\!\rangle \mathsf{Mm}_{0.9}\mathsf{f}]_{Gg} = 0.791$, *and* $[\langle\!\langle a\rangle\!\rangle \mathsf{Mm}_{0.9}\mathsf{f}]_{gg} = 0.9$. *We note that, in the course of computing the values, it turns out that using only individuals with recessive genes for mating is the best policy when we want to maximize the expected average fitness discounted with* 0.9.

On the other hand, mating with hybrids proves best if we want to minimize *the expected average fitness (with discount* 0.9*) from state GG on; for initial states* Gg *and* gg*, mating with dominant genes gives the worst expectancy, yielding the following values:* $[[\![a]\!]\mathsf{Mm}_{0.9}\mathsf{f}]_{GG} = 0.484$, $[[\![a]\!]\mathsf{Mm}_{0.9}\mathsf{f}]_{Gg} = 0.464$, *and* $[[\![a]\!]\mathsf{Mm}_{0.9}\mathsf{f}]_{gg} = 0.507$.

We observe that various levels of satisfaction and validity of MTL$_1$ formulae (and thus also the typical computational problems) can be defined analogously to Section 4.3.

The semantic definition od $\langle\!\langle a\rangle\!\rangle$ refers to the set of all stochastic policies Σ, which suggests that looking for the best policy can be quite a complex task. Is it possible to restrict the search to pure policies only? Unfortunately, it turns out that it is not the case in general. However, we conjecture that an analogous property should hold for the "state-based" fragment of MTL$_1$.

Proposition 13 *Let $\vartheta \equiv \langle\!\langle a \rangle\!\rangle\, \varphi$ be a formula of* MTL_1*. Then, the following does* not *hold:* $[\langle\!\langle a \rangle\!\rangle\, \varphi]_{\mathcal{M},q} = \sup_{s\in\sigma_{\mathcal{M}}}\{[\varphi]_{\mathcal{M}\dagger s,q}\}$*. It does not even hold for* deterministic labeled transition systems*, i.e., Markov decision processes where all transitions have probability* 1 *and all utility fluents take only classical truth values* $\top, \bot$*.*

Proof. We use the MDP from Figure 6A and the formula $\vartheta \equiv (\mathsf{M}\bigcirc \mathsf{p} \wedge \mathsf{M}\bigcirc \neg\mathsf{p})$ as a counterexample. There are two available pure strategies: $s(q_1) = \alpha$ and $s(q_1) = \beta$ (the action at q_2 is pre-determined); the instantiations $\mathcal{M} \dagger s_1$ and $\mathcal{M} \dagger s_2$ are shown in Figures 6B and 6C. For both resulting Markov chains, we have $[\mathsf{M}\bigcirc \mathsf{p} \wedge \mathsf{M}\bigcirc \neg\mathsf{p}]_{\mathcal{M}\dagger s_1,q_1} = 0$ and $[\mathsf{M}\bigcirc \mathsf{p} \wedge \mathsf{M}\bigcirc \neg\mathsf{p}]_{\mathcal{M}\dagger s_2,q_1} = 0$. In consequence, $\sup_{s\in\sigma_{\mathcal{M}}}\{[\varphi]_{\mathcal{M}\dagger s,q}\} = 0$.

Consider now the stochastic strategy $s_3(q_1,\alpha) = s_3(q_1,\beta) = 0.5$ (cf. Figure 6D). It is easy to see that $[\mathsf{M}\bigcirc \mathsf{p} \wedge \mathsf{M}\bigcirc \neg\mathsf{p}]_{\mathcal{M}\dagger s_3,q_1} = 0.5$. Thus, $[\vartheta]_{\mathcal{M},q_1} \geq 0.5$, which cannot be obtained by any pure strategy. ∎

Conjecture 14 *Let $\vartheta \equiv \langle\!\langle a \rangle\!\rangle\, \varphi$ be a formula of* MTL_1 *in which every occurrence of a temporal operator is immediately preceded by exactly one path quantifier, and every occurrence of a path quantifier is immediately preceded by exactly one strategic operator. Then:* $[\langle\!\langle a \rangle\!\rangle\, \varphi]_{\mathcal{M},q} = \sup_{s\in\sigma_{\mathcal{M}}}\{[\varphi]_{\mathcal{M}\dagger s,q}\}$.

5.2 Beyond MDP: the Multi-Agent Case

In the more general case, a system can include multiple agents/processes, interacting with each other. Here, we only briefly discuss how Markov temporal logic can be extended to handle such interaction.

On the language level, we propose to extend the strategic operator $\langle\!\langle a \rangle\!\rangle$ to a family of operators $\langle\!\langle A \rangle\!\rangle$, parameterized with groups of agents A. Intuitively $\langle\!\langle A \rangle\!\rangle\, \varphi$ refers to how much agents A can "make out of" φ by following their best joint policy. This would yield a language similar to the alternating-time temporal logic ATL* from [2], albeit with strategic operators separated from path quantifiers.

On the semantic level, multi-agent Markov decision processes [8] can be used as models. The semantics $\langle\!\langle A \rangle\!\rangle\, \varphi$ should be of course based on the maximal value of φ with respect to A's joint strategies. However, it is not entirely clear how *the other agents'* actions should be fixed in order to instantiate the MMDP to a Markov chain. One option is to assume that the opponents play a strategy that minimizes φ best. This way, operator $\langle\!\langle A \rangle\!\rangle$ would correspond to the maxmin of the two-player game where A is the (collective) maximizer, and the rest of agents fills in the role of the (collective) minimizer. Still, such a semantics would entail a very strong assumption, namely that the opponents of A must also play only *memoryless* strategies.

6 Comparison to DCTL (de Alfaro et al.)

Markov temporal logic (MTL), proposed in this paper, is in many respects similar to the "Discounted CTL" (DCTL) by de Alfaro and colleagues [10, 11]. This section lists some differences between both logics.

1. In DCTL, the set of truth values is $[0, 1]$. We keep the choice more open: it can be any continuous subset of $\mathbb{R} \cup \{-\infty, +\infty\}$.
2. MTL has more general syntax than DCTL: MTL_0 extends CTL* and MTL_1 extends the single-agent fragment of ATL*, while de Alfaro et al.'s DCTL extends only the "vanilla" CTL (despite the fact that they extend their valuation function to simple path formulae anyway).
3. E, A are true path quantifiers in our framework, in the sense that they refer to "limit properties" of paths, like the *existence* of a path with a particular (cumulative or average) utility. For aggregation of utilities via expected value, we propose a separate path operator M. In contrast, [10, 11] propose a semantics in which both E, A are based on the expected reward. In consequence, neither universal nor existential quantification on paths is expressible in DCTL for models with quantitative transition relations. One peculiar consequence of such approach is that the DCTL's $\mathsf{E}\gamma$ yields the same truth value as $\mathsf{A}\gamma$ for all Markov chains, which is not the case in our framework.

   The reason lies probably in the fact that de Alfaro and colleagues implicitly assume their path quantifiers to quantify over *actions* (or, more generally, *strategies*), while we assume that they should quantify over *transitions* (more generally, *paths*); if quantification over strategies is needed, it may be done via separate modalities like the $\langle\langle A \rangle\rangle$ of alternating-time logic ATL. Another consequence is that the semantics of path quantifiers in [10, 11] is different for qualitative and quantitative models, which is not the case in our semantics.[7]
4. MTL includes the operator $\preccurlyeq$, which can serve both as a kind of crisp material implication for fuzzy operands, and as a "defuzzification" operator that maps quantitative characteristics to classical (qualitative) descriptions.
5. The last feature allows us to define the notions of satisfaction and validity. Thus, standard problems like satisfiability and validity are properly defined in our framework.

[7] A technical remark: de Alfaro et al. call their qualitative models *labeled transition systems*, but what they consider is in fact *unlabeled* transition systems, as the evolution of the system is given by a single transition relation (i.e., there is no distinction between different actions or choices, either explicit or implicit). Thus, there is a certain discrepancy between the ultimate qualitative and quantitative models that they use: their transition systems are action-less, while Markov decision processes do include actions.

6. The "defuzzification" operator $\preccurlyeq$ enables to define strong equivalence between formulae in the sense of equality between their truth values (cf. the derived operator $\cong$). Thus, it is enough to add relevant constant utility fluents (representing particular utility values) to model-check statements like "the guaranteed average reward is at least 0.2" ($0.5 \preccurlyeq \mathsf{Am}\,\mathsf{reward}$) or "the expected maximal reward is exactly 0.5" ($\mathsf{M}\Diamond\mathsf{reward} \cong 0.5$).

7. MTL includes the full "until" operator $\mathcal{U}$, while DCTL includes only "sometime" ($\Diamond$).

8. The "always" operator is discounted in a straightforward way here. In [10, 11], it is discounted with the complement of its discount factor in order to maintain the standard relationship between "always" and "sometime".

9. We propose only the ordinary semantics for MTL (it is called the "path semantics" in [10, 11]). We believe it is more appropriate to introduce explicit fixpoint operators, rather than to define two different semantics of the same formulae.

10. In contrast to [10, 11], we do not try to capture strategic properties of the decision-making agent with temporal path quantifiers. Instead, we propose to use an explicit strategic quantifier $\langle\!\langle a \rangle\!\rangle$.

In essence: we attempt at a more *systematic* exploration of linguistic features that are offered by propositional modal logic for analysis of Markovian models of agents.

7 Conclusions

Two kinds of models are used in multi-agent systems to represent and reason about the behavior of agents/processes: quantitative and qualitative ones. In this paper, we suggest that both traditions are complementary rather than competitive. In fact, we believe that an integration of both approaches may bring a really powerful framework for dealing with multi-agent systems. To this, end, we propose *Markov temporal logic* MTL which can be seen as an extension of "Discounted CTL" from [11]. We show that the simplest version of MTL (for Markov chains) strictly extends the branching-time logic CTL*, and we discuss some fixpoint properties for a "state-based" subset of the logic. Finally, we discuss how the basic logic can be extended to address strategic abilities of agents in Markov decision processes, in a way similar to the alternating-time temporal logic ATL*.

References

[1] R. Alur, T. A. Henzinger, and O. Kupferman. Alternating-time Temporal Logic. In *Proceedings of the 38th Annual Symposium on Foundations of Computer Science (FOCS)*, pages 100–109. IEEE Computer Society Press, 1997.

[2] R. Alur, T. A. Henzinger, and O. Kupferman. Alternating-time Temporal Logic. *Journal of the ACM*, 49:672–713, 2002.

[3] A. Aziz, V. Singhal, R. K. Brayton, and A. L. Sangiovanni-Vincentelli. It usually works: The temporal logic of stochastic systems. In *Proceedings of CAV*, volume 939 of *LNCS*, pages 155–165, 1995.

[4] F. Bacchus. Probabilistic belief logics. In *Proceedings of ECAI*, pages 59–64, 1990.

[5] A. Baltag. A logic for suspicious players. *Bulletin of Economic Research*, 54(1):1–46, 2002.

[6] R. Bellman. *Dynamic Programming*. Princeton University Press, 1957.

[7] R. Bellman. A Markovian decision process. *Journal of Mathematics and Mechanics*, 6:679–684, 1957.

[8] Craig Boutilier. Sequential optimality and coordination in multiagent systems. In *Proceedings of IJCAI*, pages 478–485, 1999.

[9] E.M. Clarke and E.A. Emerson. Design and synthesis of synchronization skeletons using branching time temporal logic. In *Proceedings of Logics of Programs Workshop*, volume 131 of *Lecture Notes in Computer Science*, pages 52–71, 1981.

[10] L. de Alfaro, M. Faella, T.A. Henzinger, R. Majumdar, and M. Stoelinga. Model checking discounted temporal properties. In *Proceedings of TACAS'04*, volume 2988 of *LNCS*, pages 57–68, 2004.

[11] L. de Alfaro, M. Faella, T.A. Henzinger, R. Majumdar, and M. Stoelinga. Model checking discounted temporal properties. *Theoretical Computer Science*, 345:139–170, 2005.

[12] A. W. Drake. *Observation of a Markov Process Through a Noisy Channel*. PhD thesis, Massachusetts Institute of Technology, 1962.

[13] Steve Easterbrook and Marsha Chechik. A framework for multi-valued reasoning over inconsistent viewpoints. In *International Conference on Software Engineering*, pages 411–420, 2001.

[14] E. A. Emerson. Temporal and modal logic. In J. van Leeuwen, editor, *Handbook of Theoretical Computer Science*, volume B, pages 995–1072. Elsevier Science Publishers, 1990.

[15] E.A. Emerson and J.Y. Halpern. "sometimes" and "not never" revisited: On branching versus linear time temporal logic. *Journal of the ACM*, 33(1):151–178, 1986.

[16] P. Godefroid, M. Huth, and R. Jagadeesan. Abstraction-based model checking using modal transition systems. In *Proceedings of CONCUR*, volume 2154 of *LNCS*, pages 426–440, 2001.

[17] C. M. Grinstead and J. L. Snell. *Introduction to Probability*. Amer Mathematical Society, 1997.

[18] P. Hajek. *Metamathematics of Fuzzy Logic*. Kluwer, 1998.

[19] J. Y. Halpern. A logical approach to reasoning about uncertainty: a tutorial. In X. Arrazola, K. Korta, and F. J. Pelletier, editors, *Discourse, Interaction, and Communication*, pages 141–155. Kluwer, 1998.

[20] H. Hansson and B. Jonsson. A logic for reasoning about time and reliability. *Formal Aspects of Computing*, 6(5):512–535, 1994.

[21] P. Harrenstein, W. van der Hoek, J-J. Meijer, and C. Witteveen. Subgame-perfect Nash equilibria in dynamic logic. In M. Pauly and A. Baltag, editors, *Proceedings of the ILLC Workshop on Logic and Games*, pages 29–30. University of Amsterdam, 2002. Tech. Report PP-1999-25.

[22] J. Hintikka. *Logic, Language Games and Information*. Clarendon Press : Oxford, 1973.

[23] R. A. Howard. *Dynamic Programming and Markov Processes*. MIT Press, 1960.

[24] J. G. Kemeny, L. J. Snell, and A. W. Knapp. *Denumerable Markov Chains*. Van Nostrand, 1966.

[25] G. J. Klir and T. A. Folger. *Fuzzy Sets, Uncertainty and Information*. Englewood Cliffs: Prentice Hall, 1988.

[26] B. Konikowska and W. Penczek. Model checking for multi-valued CTL*. In M. Fitting and E. Orlowska, editors, *Multi-Valued Logics*. 1998.

[27] A. Lluch-Lafuente and U. Montanari. Quantitative μ-calculus and CTL based on constraint semirings. *Electr. Notes Theor. Comput. Sci.*, 112:37–59, 2005.

[28] K. Lorenz and P. Lorenzen. *Dialogische Logik*. Darmstadt, 1978.

[29] A. Markov. Rasprostranenie zakona bol'shih chisel na velichiny, zavisyaschie drug ot druga. *Izvestiya Fiziko-matematicheskogo obschestva pri Kazanskom universitete*, 2(15):135–156, 1906.

[30] F. Moller and A. Rabinovich. On the expressive power of CTL*. In *Proceedings of LICS'99*, pages 360–369, 1999.

[31] N. J. Nilsson. Probabilistic logic. *Artificial Intelligence*, 28(1):71–87, 1986.

[32] E.H. Ruspini, J. Lowrance, and T. Strat. Understanding evidential reasoning. *Artificial Intelligence*, 6(3):401–424, 1992.

[33] R. S. Sutton and A. G. Barto. *Reinforcement Learning: An Introduction*. MIT Press, 1998.

[34] J. van Benthem and F. Liu. Dynamic logic of preference upgrade. *Journal of Applied Non-Classical Logic*, 17(2), 2007.

[35] W. van der Hoek. *Modalities for Reasoning about Knowledge and Quantities*. PhD thesis, ILLC Amsterdam, 1992.

[36] W. van der Hoek, W. Jamroga, and M. Wooldridge. A logic for strategic reasoning. In *Proceedings of AAMAS'05*, pages 157–164, 2005.

[37] D. Williams. *Probability with Martingales*. Cambridge University Press, 1991.

[38] P. Xuan, V. R. Lesser, and S. Zilberstein. Communication in multi-agent markov decision processes. In *Proceedings of ICMAS*, pages 467–468, 2000.

[39] L.A. Zadeh. Fuzzy sets. *Information and Control*, 8(3):338–353, 1965.